A Review on Architecture of Deep Learning

Dr. V. Ramesh
Assistant Professor
SCSVMV University, Kanchipuram, Tamil Nadu, India

Abstract— In recent years we have seen an amazing improvement in applications using Deep learning. It started with speech recognition then moved on to computer vision, object recognition and natural language processing. Deep learning constitutes a recent, modern technique for image processing and data analysis, with promising results and large potential. Deep learning are machine learning algorithms based on learning multiple level of abstraction. Deep learning allows computational models that are composed of multiple processing layers to learn representations of data with multiple levels of abstraction. These methods have dramatically improved the state-of-the-art in speech recognition, visual object recognition, object detection and many other domains such as drug discovery and genomics. Deep learning discovers intricate structure in large data sets by using the backpropagation algorithm to indicate how a machine should change its internal parameters that are used to compute the representation in each layer from the representation in the previous layer. Deep convolutional nets have brought about breakthroughs in processing images, video, speech and audio, whereas recurrent nets have shone light on sequential data such as text and speech. In this study we have done a survey to provide a general overview on the novel concept and the ever-increasing advantages and popularity of deep learning.

Keywords: Deep Learning, Classification, Regression

I. Introduction

It is expected that global population will reached 9 billion in 2050.it means we require more agricultural production in order to meet food demands. Otherwise we will be suffering from food security problems. So we have to increase our per unit area production. In agriculture we emphasis on yield. For increasing yield, our farm enterprises require new and innovative technologies to face and overcome these difficult challenges. So we are using different methodologies, technologies, and different processes for higher yield per unit area. In methodological approaches we can use different simulation models, but nowadays deep neural network or deep learning is also used in modern countries. Deep learning is a discipline of Computer Science that deals with giving ability to the machines that it seems to behave like, that it has human intelligence. There is a lot of scope of deep learning in many fields of life like in Agriculture that it can monitors crop conditions, like water scarcity conditions, plant population in field and soil moisture content etc. Deep learning is working in almost all disciplines of agriculture. In Irrigation it can control irrigation water in the field we can optimize the use of water because it have automatic irrigation systems in it that take weather conditions and predicts amount of water to be applied. Non chemical weed control is used in discriminating between weeds and crop seedling. It is an important step towards control of weeds by nonchemical way. Drone technology is another adoption of Artificial intelligence. Drones can be used to provide detailed mapping of crops in the fields. They are also capable of delivering customized fertilizers pesticides, insecticides based upon the requirement of each crops.

Artificial intelligence and deep learning algorithms has been found useful in almost every field of work and study. It being deployed in every field makes it the next big thing and breakthrough for a smart future. In India, especially in Punjab increasing agriculture output to meet ever increasing population's demand is one of the major issues being faced today. Agricultural experts would tell that factors like climate, soil, rain along with other factors affect the output of a crop. Farmers feel immense pressure in such situations and lack of knowledge to counter the problems faced and modern techniques of farming only add further to the problem. There have been various sorts of efforts to implement deep learning and check the innumerable effect it has on different aspects like production time, assistance, output etc. We have seen field of medical science specifically benefitting from the implementation of deep neural networks. Researchers are constantly working of technologies like machine learning, deep neural network to aid artificial intelligence as a product for human use.

II. Related Work

Although deep learning has historical roots going back decades (Schmidhuber, 2015), it attracted relatively little notice until just over five years ago. Virtually everything changed in 2012, with the publication of a series of highly influential papers such as Krizhevsky, Sutskever and Hinton's 2012 ImageNet Classification with Deep Convolutional Neural Networks (Krizhevsky, Sutskever, & Hinton, 2012), which achieved state-of-the-art results on the object recognition challenge known as ImageNet (Deng et al.,). Other labs were already working on similar work (Ceresan, Meier, Masci, & Schmidhuber, 2012). Before the year was out, deep learning made the front page of The New York Times, and it rapidly became the 2 best known technique in artificial intelligence, by a wide margin. If the general idea of training neural networks with multiple layers was not new, it was, in part because of increases in computational power and data, the first time that deep learning truly became practical.

Deep learning has since yielded numerous state of the art results, in domains such as speech recognition, image recognition, and language translation and plays a role in a wide swath of current AI applications. Corporations have invested billions of dollars fighting for deep learning talent. One prominent deep learning advocate, Andrew Ng, has gone so far to suggest that "If a typical person can do a mental task with less than one second of thought, we can probably automate it using AI either now or in the near future."

III. Deep Learning Architectures

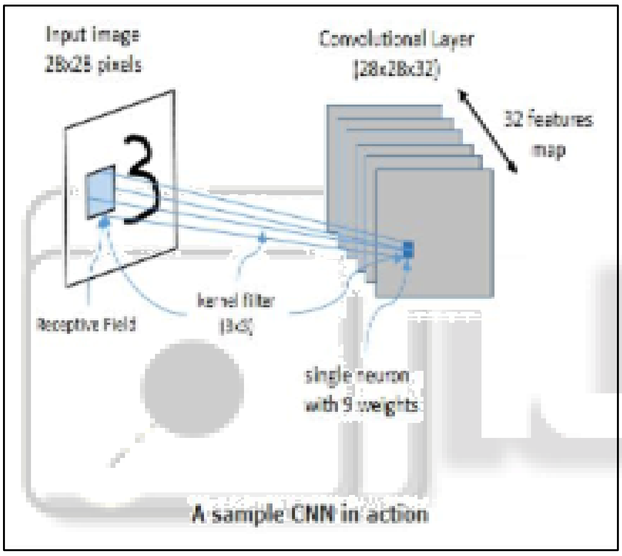
There are huge numbers of variants of deep architectures. Most of them are branches of some original parent architecture.

A. Convolutional Neural Networks

Convolutional Neural Networks, or CNNs in short, are the popular choice of neural networks for different Computer Vision tasks such as image recognition. The name 'convolution' is derived from a mathematical operation involving the convolution of different functions. There are 4 primary steps or stages in designing a CNN:

1) Convolution: The input signal is received at this stage
2) Subsampling: Inputs received from the convolution layer are smoothened to reduce the sensitivity of the filters to noise or any other variation
3) Activation: This layer controls how the signal flows from one layer to the other, similar to the neurons in our brain
4) Fully connected: In this stage, all the layers of the network are connected with every neuron from a preceding layer to the neurons from the subsequent layer.

Here is an in-depth look at the CNN Architecture and its working, as explained by the popular AI Researcher Giancarlo Zaccone.

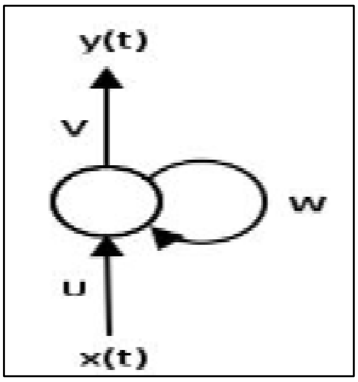


A sample CNN in action

B. Recurrent Neural Networks

Recurrent Neural Networks (RNNs) have been very popular in areas where the sequence in which the information is presented is crucial. As a result, they find a lot applications in real-world domains such as natural language processing, speech synthesis and machine translation. RNNs are called 'recurrent' mainly because a uniform task is performed for every single element of a sequence, with the output dependent on the previous computations as well. Think of these networks as having a memory, where every calculated information is captured, stored and utilized to calculate the final outcome. Over the years, quite a few varieties of RNNs have been researched and developed

1) Bidirectional RNN – The output in this type of RNN depends not only on the past but also the future outcomes
2) Deep RNN – In this type of RNN, there are multiple layers present per step, allowing for a greater rate of learning and more accuracy.
3) RNNs can be used to build industry-standard chatbots that can be used to interact with customers on websites. Given a sequence of signals from an audio wave, RNNs can also be used to predict a correct sequence of phonetic segments with a given probability.



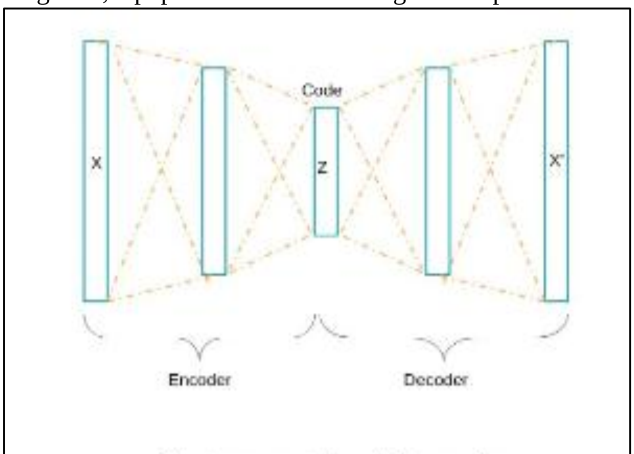
C. Autoencoders

Autoencoders apply the principle of backpropagation in an unsupervised environment. Autoencoders, interestingly, have a close resemblance to PCA (Principal Component Analysis) except that they are more flexible. Some of the popular applications of Autoencoders is anomaly detection – for example detecting fraud in financial transactions in banks. Basically, the core task of autoencoders is to identify and determine what constitutes regular, normal data and then identify the outliers or anomalies.

Autoencoders usually represent data through multiple hidden layers such that the output signal is as close to the input signal. There are 4 major types of autoencoders being used today:

1) Vanilla autoencoder: The simplest form of autoencoders there is, i.e. a neural net with one hidden layer
2) Multilayer autoencoder: When one hidden layer is not enough, an autoencoder can be extended to include more hidden layers
3) Convolutional autoencoder: in this type, convolutions are used in the autoencoders instead of fully-connected layers
4) Regularized autoencoder: This type of autoencoders use a special loss function that enables the model to have properties beyond the basic ability to copy a given input to the output.

This article demonstrates training an autoencoder using H20, a popular machine learning and AI platform.

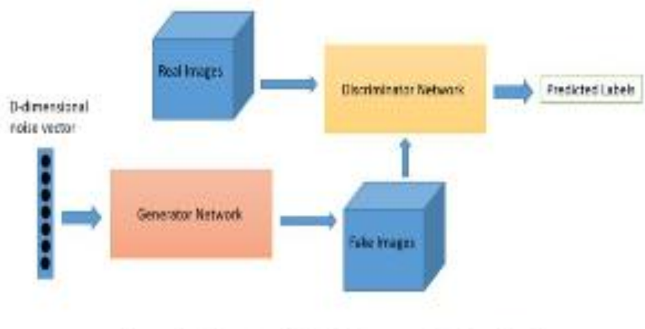


A basic representation of Autoencoder

D. Generative Adversarial Networks

The basic premise of Generative Adversarial Networks (GANs) is the training of two deep learning models simultaneously. These deep learning networks basically compete with each other – one model that tries to generate new instances or examples is called as the generator. The other model that tries to classify if a particular instance originates from the training data or from the generator is called as the discriminator.

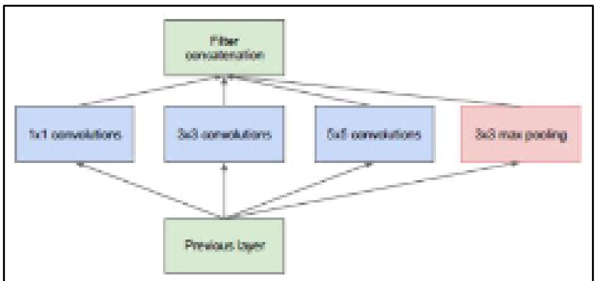
GANs, a breakthrough recently in the field of deep learning was a concept put forth by the popular deep learning expert Ian Goodfellow in 2014. It finds large and important applications in Computer Vision, especially image generation.



General architecture of GAN (Source: deeplearning4j)

E. *GoogleNet*

GoogleNet (or Inception Network) is a class of architecture designed by researchers at Google. GoogleNet was the winner of ImageNet 2014, where it proved to be a powerful model. In this architecture, along with going deeper (it contains 22 layers in comparison to VGG which had 19 layers), the researchers also made a novel approach called the Inception module.

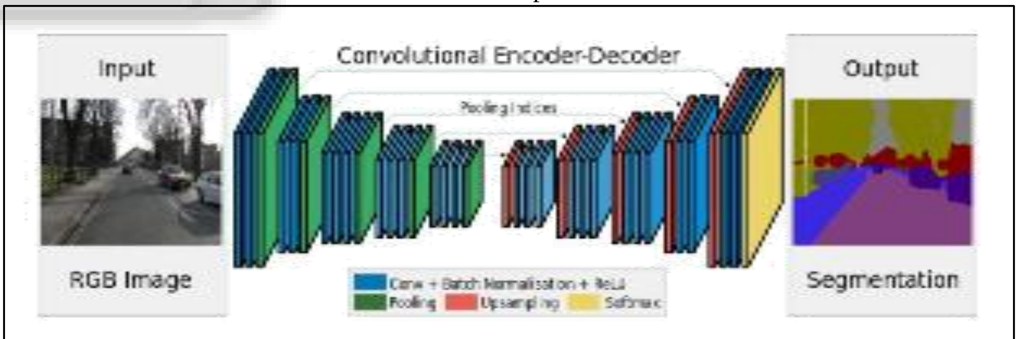


(a) Inception module, naïve version

As seen above, it is a drastic change from the sequential architectures which we saw previously. In a single layer, multiple types of "feature extractors" are present. This indirectly helps the network perform better, as the network at training itself has many options to choose from when solving the task. It can either choose to convolve the input, or to pool it directly. The final architecture contains multiple of these inception modules stacked one over the other. Even the training is slightly different in GoogleNet, as most of the topmost layers have their own output layer. This nuance helps the model converge faster, as there is a joint training as well as parallel training for the layers itself.

F. *SegNet*

SegNet is a deep learning architecture applied to solve image segmentation problem. It consists of sequence of processing layers (encoders) followed by a corresponding set of decoders for a pixelwise classification. Below image summarizes the working of SegNet. One key feature of SegNet is that it retains high frequency details in segmented image as the pooling indices of encoder network is connected to pooling indices of decoder networks. In short, the information transfer is direct instead of convolving them. SegNet is one the the best model to use when dealing with image segmentation problems.



IV. CONCLUSION

Deep learning is indeed a fast growing application of machine learning. During the current era and in a future, deep learning can result into a useful security tool due to the facial recognition and speech recognition combined. Besides this, digital image processing is a research field that can be applied in multiple areas. For this reason and having proved a true optimization, deep learning is a contemporary and exciting subject of advancement in artificial intelligence.

REFERENCES

[1] R. Vargas, A. Mosavi. "DEEP LEARNING: A REVIEW." Advances in Intelligent Systems and Computing (2017).
[2] Y. Bengio, "Learning deep architectures for AI," Foundations and trends® in Machine Learning, vol. 2, pp. 1-127, 2009.
[3] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, pp. 436-444, 2015
[4] Verma Kunal, Pabbi Dinesh, ''Agriculture Advancement Using Artificial Intelligence'', International conference on recent innovations in science, technology, management and enironment, 2016.